Investigating parameter risk for Solvency II and ICAS

Internal models are incentivised under Solvency II by the potential to lower capital requirements, but are subject to a number of risks. Parameter risk is one that has got scant attention. This article demonstrates that the primary source of parameter risk for annuity providers is the direction of future trends, and argues that insurers should hold additional capital accordingly.

BY STEPHEN RICHARDS

> *"The longevity shock to be applied is a (permanent) 25% decrease in mortality rates for each age."*
> Ceiops (2008) TS.XI.C.6

> *"Such parameters shall be calibrated on the basis of the internal data of the undertaking concerned, or of data which is directly relevant for the operations of that undertaking using standardised methods."*
> Article 104, Commission of the European Communities (2008)

> *"A firm should identify and justify key assumptions and document the sensitivity of the results to these assumptions and its approach to dealing with parameter uncertainty and fluctuations."*
> Principle 2.1, FSA (2006)

THE ADVENT OF SOLVENCY II creates a big incentive to upgrade internal modelling capabilities. For example, the QIS4 (Quantitative Impact Study 4) rules for annuities and pensions can lead to large increases in reserves at older ages. In order to demonstrate to regulators that such capital increases are unnecessary, a life office must use appropriate internal models. Ideally, these models are parameterised using each office's own experience data: the QIS4 rules on their own are too crude to use for internal capital allocation and risk management.

There are a variety of issues associated with any kind of modelling. The first is model risk, namely the fact that it is impossible to be sure if the model structure you are using is the appropriate one. For example, different models can indeed be compared in their overall fit to the data, say by using the AIC (Akaike, 1987) or some other objective criterion. However, just because a model fits the existing data well does not automatically mean it will be the best-fitting in the future. Model risk as applied to mortality projections for annuities and pensions is addressed in Richards and Currie (2009), while Cairns (2004) discusses model risk in application to interest rates.

A second issue is basis risk, which arises when you explain a random process in terms of another process, but in doing so, make an implicit assumption about the relationship between those processes and the persistency of that relationship. In mortality modelling, a larger dataset might be used to define much of the characteristics of a model. For example, few life offices have sufficient historical data on mortality rates for pensioners or annuitants, and it is common to fall back on population statistics to model mortality trends. However, a portfolio of private pensioners is a very select subset of the wider population, and there is the risk that the trends exhibited in the two data sets are not the same.

The third issue, and the focus of this article, is parameter risk. Even if you have the correct model, and even if you have a statistically credible set of the portfolio's own experience data, parameters are still subject to estimation error. It is of interest to regulators and investors what impact this estimation error could have financially. In this article we will illustrate the points made with reference to a portfolio of pension annuitants in the United Kingdom. However, the points made are general and will apply to other classes of business and other national territories.

Parameter risk

Life offices can now fit quite sophisticated survival models to their demographic experience data. Richards (2008) demonstrated how age, gender, pension size, postcode, cohort, select period and time trend could all be included in a model for pensioner mortality. For example, a robust model for pensioner mortality was given by Perks (1932), which can be extended to include a simple time trend:

$$\mu_{x,y} = \frac{\exp(\alpha + \beta x + \delta y)}{1 + \exp(\alpha + \beta x + \delta y)}$$

where $\mu_{x,y}$ denotes the instantaneous force of mortality at exact age x and y denotes calendar time in years measured from 1 January 2000. This latter parameterisation is useful for keeping the parameters well-scaled, and means that 1 January 1998 would be represented as –2, whereas 1 July 2005 would be +5.5. "Well-scaled" here means keeping the various parameter values working on broadly similar scales. For example, under this parameterisation values for α tend to fall into the range (–14, –10), while β tends to fall into the range (0.09, 0.14) and δ usually lies in the range (–0.05, 0). If we multiplied δ by the calendar year directly, then the range of values would be at least two orders of magnitude smaller, and fitting algorithms often run into difficulties when parameters have very different scales. There is nothing to stop us measuring time from 1 July 2005, for example, or any other contemporary date for that matter, but using 1 January 2000 is perhaps the most convenient in practice.

The parameter α is called the intercept, since it approximately determines the log force of mortality at age zero and time zero; that is, the point at which mortality crosses or intercepts the joint age and time axis. β measures the age-related change in mortality, and represents broadly the change in mortality rate due to an increase of one year of age. δ measures the time trend in mortality, and is a deliberately crude-but-simple model for time-based patterns. δ represents broadly the change in overall mortality level due to a one-year change in calendar time. In a regression model for pensioner mortality we can allow different sub-groups to have different values of α, β and δ, as shown in Richards (2008). We could vary δ by sub-groups, for example by fitting an interaction with gender to allow males and females to have different constant rates of improvement. The maximum-likelihood estimates for a simple model with age, gender and time are shown in Table 1. Females are taken as the baseline in the model, so the gender effect of 1.80965 is the addition to the intercept for males, while the age:gender interaction of –0.0174623 is the addition to the age coefficient for males. More formally, this model would be expressed as follows in relation to Equation 1:

$$\alpha = \alpha_0 + \alpha_{male} \times z_{male}$$
$$\beta = \beta_0 + \beta_{male} \times z_{male}$$

where z_{male} is an indicator variable taking the value 0 for females and 1 for males.

Table 1: Parameter estimates for mortality model Age*Gender+Time

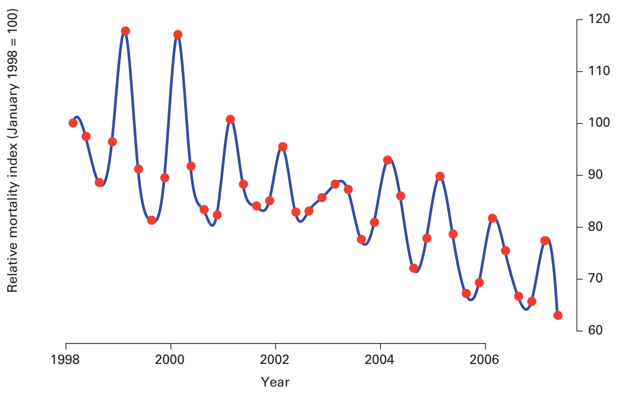
	Estimate	Standard error	Z-value	Coefficient of variation
Age (β_0)	0.131207	0.001	125.08	0.80%
Gender.M (α_{male})	1.80965	0.098	18.46	5.42%
Gender.M:Age (β_{male})	–0.0174623	0.0013	–13.96	7.16%
Intercept (α_0)	–13.586	0.0827	–164.27	0.61%
Time (δ)	–0.03397	0.0017	–19.95	5.01%

Source: Longevitas Ltd. Perks law for force of mortality fitted to mortality-experience data of annuitants aged between 60 and 95 over 1998{2006. The Z-values and coefficients of variation are calculated using the unrounded standard errors.

The null hypothesis is that the underlying value for each parameter is zero. We have fitted these parameters such that they jointly maximise the log-likelihood function, and the estimated standard errors are derived from inverting the matrix for second derivatives. As maximum-likelihood estimates (MLEs), their distribution is asymptotically normal. To test the null hypothesis, we compute the Z-value for comparing against the N(0, 1) distribution, by dividing the parameter estimate by its approximate standard error. The Z-values in Table 1 are well outside the range (–2, +2), so we know that each of the parameters is significantly different from zero. Of particular interest here is the time-trend parameter, which shows mortality rates falling by around 3.3% per annum continuously over the period 1998–2006 (3.3% = 100% – exp(–0.03397)). Figure 1 shows this using a different model, which also demonstrates clear annual and seasonal fluctuations. However, the assumption of a constant rate of improvement is a workable starting assumption for a long-term liability like pensions, which could be refined by fitting interactions with other risk factors, such as gender.

While all the parameter estimates in Table 1 are highly significant, they have different levels of relative certainty over the fitted values. The coefficient of variation is the standard error expressed as a percentage of the parameter estimate to get an idea of how narrow or wide the confidence intervals for each estimate actually is in relation to the parameter itself. This is shown in the final column of Table 1, where it should be

Figure 1. Mortality levels by season for a large annuity portfolio



The solid red dots show the effect of season on overall mortality levels for a portfolio of annuitants aged between 60 and 95. The seasonal effects are indexed with winter 1998 set to 100. Results from fitted Cox survival model with age, gender and season as rating factors, where the definition of season follows that used by the ONS: winter (Nov–Jan), spring (Feb–Apr), summer (May–Jul) and autumn (Aug–Oct). Source: Longevitas Ltd.

noted that the coefficient of variation is also simply the absolute value of the inverse of the Z-value. We can see that there is far more certainty over the intercept and the rate of ageing (coefficient of variation under 1%) than there is for the time-trend or the gender-related parameters.

To get our baseline of stochastic and concentration risk, we take the maximum-likelihood parameter estimates in Table 1 and simulate the lifetime of every pensioner in the portfolio. This gives us a single portfolio simulation, which we can use to calculate the present value of pension payments made. We can then repeat this portfolio simulation 10,000 times and look at the distribution of present values. Taking the parameter estimates in Table 1 as fixed, our results show the variation due to stochastic risk and concentration risk. The row marked “No perturbation” in Table 2 shows the relative percentage difference of various scenarios compared to the median. For example, by sorting the costs of the 10,000 scenarios, we can look at the 9950th most expensive to get the 99.5% stress scenario in run-off of 0.48% extra cost relative to the median. The most expensive of the 10,000 run-offs produced a present value of 0.70% higher than the median. Note that the figures in Table 2 are underestimates of variability, since our model assumes a constant rate of improvement, whereas Figure 1 shows that there is considerable inter- and intra-year fluctuation.

However, the line marked “No perturbation” in Table 2 assumes that the parameters are precisely known. This is not the case, of course, as there is uncertainty over the underlying parameter estimates, as shown in Table 1. Assuming the model is correct, then the MLEs in Table 1 are asymptotically normal, with the mean being the true parameter in each case. Since this portfolio is very large, we can substitute the estimate and standard error calculated in Table 1 for the true unknown underlying values of mean and standard error. This gives us the opportunity to examine the impact of parameter uncertainty: for each portfolio simulation we can perturb a parameter using a pseudo-random N(0, 1) variate and the estimated standard error. We can then re-fit the model to get maximum-likelihood estimates for the other parameters, subject to the perturbed value being fixed at its new value. With this new perturbed model we can then conduct a simulated run-off of the entire portfolio covering both binomial risk and concentration risk, as before. By repeating this 10,000 times, with the given parameter being perturbed once per portfolio simulation, we can explore the financial impact of uncertainty over each parameter. The results of this are shown for each parameter in Table 2.

The first line in Table 2 shows the combined stochastic and run-off risks for the portfolio; that is, just using the maximum-likelihood estimates in Table 1 without any perturbation. The relative impact of these two risks is small, as the portfolio is large. The next four rows show the impact of perturbations in the parameters for the intercept, age, gender, and the age:gender interaction. Comparing these with the no-perturbation run-off results, we can see that uncertainty over these parameters is adding negligible further risk at the 99.5% level. However, it is the results for the impact of uncertainty over the time trend which really stand out, requiring four times the extra capital compared to the uncertainty over the other parameters. Among the 10,000 perturbations the time-trend parameter δ ranged from an implied improvement rate of 2.80% per annum to 4.02%.

The model here is admittedly very simple, and could benefit from including many more of the available rating factors. The assumption of a constant rate of improvement is crude, since mortality improvements also vary by (for example) age and gender. However, Table 2 suggests that uncertainty over future improvements is the dominant source of mortality-related financial uncertainty for this large portfolio. Before passing some important comments on data preparation, we briefly consider extending the model to include year of birth, or cohort, as a risk factor. The parameters for this model are not shown for brevity, but Table 3 shows the revised table of financial impacts due to parameter uncertainty in the same manner as Table 2. Table 3 confirms that uncertainty over the time trend is still the dominant parameter financially. However, Table 3 also shows a lower relative financial impact compared to Table 2. The suggestion is that a richer model with more risk factors can reduce the capital requirements for parameter uncertainty.

Note that in all such models we are dealing with dependent parameters. This means that a given parameter cannot be interpreted independently of the others in the model, since changing the value of one will change the maximum-likelihood estimate of the others. It also means that similar-looking parameters in different models are not strictly comparable. For example, due to the addition of the cohort parameters in Table 3, the parameters Age, Gender and Time are no longer performing the exact same role as their equivalently named counterparts in Table 2. A corollary of this is that it is not

Table 2: Variation around median run-off cost for a large annuity portfolio, both with and without parameter perturbation

Parameter perturbed	min	0.5%	1%	5%	95%	99%	99.5%	max
No perturbation	-0.74%	-0.48%	-0.44%	-0.31%	0.31%	0.43%	0.48%	0.70%
Intercept	-0.80%	-0.57%	-0.50%	-0.35%	0.36%	0.50%	0.54%	0.79%
Age	-0.75%	-0.53%	-0.47%	-0.33%	0.33%	0.47%	0.52%	0.69%
Gender	-0.74%	-0.50%	-0.44%	-0.31%	0.32%	0.45%	0.48%	0.64%
Age:Gender	-0.65%	-0.47%	-0.43%	-0.32%	0.31%	0.44%	0.49%	0.74%
Time	-3.17%	-2.24%	-2.02%	-1.40%	1.36%	1.94%	2.15%	3.05%

Source: Longevitas Ltd. 10,000 simulations of portfolio of 324,261 annuitants aged 60–95, with suitably standardised perturbations for a given parameter prior to each run-off simulation. Whole-life annuities valued continuously using portfolio’s own mortality model and discounted at 5% interest p.a.. Prior to 2008 this would have been regarded as a sensible long-term rate at which to discount. At the time of writing during the credit crisis, however, this interest rate would be regarded as high: the Bank of England base rate is 0.5%. We have stuck with 5% interest, but the range of variation shown in this table would be wider still at current lower rates of interest.

Table 3: Variation around median run-off cost for a large annuity portfolio, both with and without parameter perturbation

Parameter perturbed	min	0.5%	1%	5%	95%	99%	99.5%	max
No perturbation	-0.70%	-0.48%	-0.43%	-0.31%	0.31%	0.44%	0.48%	0.74%
Intercept	-0.77%	-0.52%	-0.48%	-0.35%	0.34%	0.48%	0.55%	0.77%
Age	-0.71%	-0.52%	-0.48%	-0.33%	0.34%	0.47%	0.53%	0.76%
Gender	-0.64%	-0.49%	-0.45%	-0.32%	0.31%	0.44%	0.50%	0.74%
Age:Gender	-0.66%	-0.47%	-0.44%	-0.32%	0.31%	0.44%	0.49%	0.74%
Time	-2.44%	-1.67%	-1.50%	-1.06%	1.07%	1.55%	1.72%	2.52%
Cohort.2	-0.64%	-0.48%	-0.44%	-0.31%	0.31%	0.44%	0.48%	0.74%
Cohort.3	-0.64%	-0.49%	-0.45%	-0.32%	0.31%	0.44%	0.49%	0.74%
Cohort.4	-0.64%	-0.50%	-0.46%	-0.32%	0.32%	0.45%	0.51%	0.74%

Source: Longevitas Ltd. 10,000 simulations of portfolio of 324,261 annuitants aged 60–95, with suitably standardised perturbations for a given parameter prior to each run-off simulation. Whole-life annuities valued continuously using portfolio's own mortality model and discounted at 5% interest per annum.

always straightforward to disentangle model risk and parameter risk.

A final important point to note is that this procedure is based on a statistical model; that is, there is a clearly structured probability framework which has been parameterised using recent experience data. What this model cannot do is predict or measure the risk of some kind of radical medical breakthrough which will radically improve longevity. There are numerous examples of such breakthroughs in living memory: the discovery and introduction of antibiotics in the 1940s, and the large reductions in cardiovascular mortality courtesy of statins and ACE inhibitors in more recent times. Since our statistical model knows nothing of the the likelihood of similar future discoveries, the variation figures in Tables 2 and 3 should be viewed as a minimum bound for the likely higher range in reality.

Data preparation

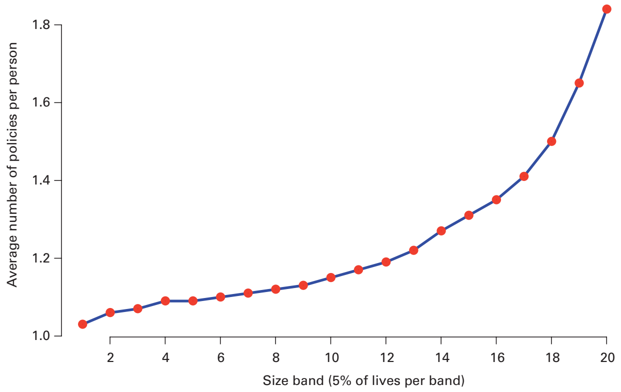
It is important to be clear about the data one uses in simulation work. We have used a data set of pension annuities in payment. The portfolio was used to simulate the lifetimes of the individuals receiving the benefits. We have simulated the lifetimes of the current holders only, assuming that benefits terminated on death of the current life. For this article we have ignored possible continuation of benefits to surviving spouses.

Prior to fitting any models or carrying out the simulations we deduplicated the portfolios using the procedure outlined in Richards (2008); that is, pensions or annuities paid to the same person were identified and aggregated. Such deduplication is essential for model-fitting due to the independence assumption for statistical models. However, deduplication is also necessary for the simulation work carried out in run-off assessments. Simulating without deduplication would be doubly misleading: first, it is common for people to have more than one annuity, so the mortality experience for annuities is not independent; second, wealthier people have a greater tendency to have multiple annuities, as shown in Figure 2, so a correct picture of the financial volatility can only be obtained by combining pension and annuity records for the same person. Failure to deduplicate prior to simulation would give a falsely comforting picture of the stochastic risk in run-off, and it would likely also underestimate the concentration risk due to large aggregated pensions. More details on the importance of deduplication are given in Richards and Currie (2009), while details on the algorithms used for deduplication are given in Richards (2008).

A key feature of data sets involving private benefits is concentration risk, which arises when benefits are unequal and liabilities are therefore concentrated in a relatively small proportion of lives. For example, Richards (2008) analysed the annual pension amounts paid to 777,111 distinct lives receiving either a pension annuity or a pension from an employer-sponsored defined-benefit scheme. Half of all the defined-benefit pensions were paid to 11.7% of the pensioners, while half of all the annuity payments went to just 7.8% of the annuitants. This concentration of liabilities is itself a form of basis risk: a model of mortality risk will be dominated by the majority of lives who constitute a minority of the liabilities.

The law of large numbers means that bigger portfolios experience smaller variation in run-off costs. This is variously called stochastic risk

Figure 2. Average number of policies by income band



Average number of policies per person in each of equal-sized membership bands ordered by total annual annuity income. Source: Richards and Currie (2009). Band 1 is the 5% of lives with smallest annual pensions, through to band 20 which is the 5% of lives with the largest annual pensions. Data taken from the pension annuity portfolio used in Richards (2008).

Table 4: Extra cost of 99.5% most expensive simulation relative to median run-off cost

Portfolio type	Number of lives	Average age (years)	Extra cost at 99.5% level assuming benefits are:	
			(a) equal	(b) actual
Pension annuitants	207,190	72.5	0.20%	0.50%
Pension annuitants	15,429	67.3	0.63%	1.07%
Pensioners in occupational scheme	2,268	67.2	2.02%	4.52%

Source: Richards and Currie (2009). Results of 10,000 simulations of portfolio of male lives aged between 40 and 90. Temporary annuities to age 90 valued continuously at 5 per cent interest per annum. Average age is weighted by annual pension amount. Population mortality levels without projections.

or binomial risk, and it applies even if you have the correct model with the correct parameters. This is shown in Table 4, where the extra cost of the 99.5% most expensive simulation relative to the median increases as the portfolio size decreases. Eventually a portfolio's size can become small enough such that this stochastic or binomial risk is the single most important risk for a small pension scheme. Table 4 also shows that increasing the size of a portfolio will result in some capital benefits: larger portfolios require less capital in respect of stochastic risk than smaller ones.

However, concentration of liabilities acts against the law of large numbers: for a given number of lives, the more unequal the benefits in a portfolio the greater the volatility in run-off costs. This is also shown in Table 4 – in each case the inequality in benefit size has increased the extra cost for the 99.5% most expensive run-off simulation. This issue is discussed in more detail in Richards and Jones (2004), where the conclusion reached was that pension schemes with fewer than 50 members should consider very carefully whether self-insurance of pensions was appropriate.

Interestingly, concentration risk is also inextricably linked to the problem of duplicates. Table 5 shows that a nearly a third of pensions are in respect of individuals with multiple policies. This reinforces the message that deduplication is essential in forming an accurate picture of liabilities for this annuity portfolio.

The primary purpose of run-off simulations using actual benefits is to establish a baseline against which to measure parameter risk. Note that we are simulating the entire portfolio, rather than relying on model points. Using model points is inappropriate when measuring stochastic risk, since the model points represent a smaller number and will overstate the level of stochastic risk in run-off.

Table 5: Proportion of portfolio by number of policies held

Number of policies per individual	Proportion of total:	
	(a) by lives	(b) by amounts
1	83.3%	67.6%
2	12.6%	20.3%
3	2.6%	6.4%
4	0.8%	2.6%
5	0.3%	1.3%
6	0.2%	0.6%
7	0.1%	0.4%
8	0.1%	0.3%
9	0%	0.1%
10	0%	0.1%
11	0%	0.1%
12	0%	0.1%
Total	100%	99.9%

Source: Richards and Currie (2009). The column for amounts does not quite add up to 100% because there are some individuals with even more than 12 policies (up to a maximum of 31 for this portfolio).

Conclusions

Insurers who rely on projections based on other data sets should consider holding extra capital margin for the basis risk. By building a model based on a portfolio's actual experience, the actuary can justifiably claim that basis risk is low and that reserving margins held in respect of basis risk can be released. This same model can be used to explore parameter risk as an internal model for Solvency II purposes. In the case of the pension annuity portfolio used in this article, parameter risk was found to be low, with the uncertainty over the pattern of future trends emerging as the most financially important parameter of all. We also showed how enriching the basic model with more risk factors reduced the impact of parameter uncertainty. Increasing the number of valid risk factors in a model can lower the capital requirements for parameter risk.

Stephen Richards is director of Richards Consulting
4 Caledonian Place, Edinburgh, EH11 2AS. Tel: +44 (0)131 315 4470
www.richardsconsulting.co.uk

References

Akaike, H. (1987)
Factor analysis and AIC
Psychometrica, 52, 317–333

CEIOPS (Committee of European Insurance and Occupational Pensions Supervisors) (2007)
QIS4 Technical Specifications, December 2007

Commission of the European Communities (2008)
Amended proposal for a Directive of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance

Financial Services Authority (2006)
ISG Expert Groups: Draft ICA Principles and Guidance
Presented for discussion on 5 June 2006, http://www.fsa.gov.uk/pubs/international/isgdraftica.pdf

Longevitas Development Team (2008)
Longevitas v2.4 user guide
http://www.longevitas.co.uk

Perks, W. (1932)
On some experiments in the graduation of mortality statistics
Journal of the Institute of Actuaries, 63, 12–40

R Development Core Team (2007)
R: A language and environment for statistical computing
R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.r-project.org

Richards, S. J. and Jones, G. L. (2004)
Financial aspects of longevity risk
Staple Inn Actuarial Society, London

Richards, S. J. (2008)
Applying survival models to pensioner mortality data
British Actuarial Journal (Volume 14)

Richards, S. J. and Currie I. D. (2009)
Assessing longevity risk and annuity pricing with the Lee-Carter model
British Actuarial Journal (to appear)